Original Article:

Article Title:

Improving the diagnostic criteria for primary liver graft non-function in adults utilizing standard and transportable laboratory parameters. An outcome based analysis.

Authors:

Mohammad A. B. Al-Freah *, Mark JW McPhail *, Elena Dionigi, Matthew R. Foxton, Georg Auzinger, Mohamed Rela, Julia A Wendon, John G O'Grady, Michael A Heneghan, Nigel D Heaton and William Bernal.

*Both authors contributed equally to this article.

Authors Affiliation:

Institute of Liver studies, King's College Hospital, London, United Kingdom.

Address for Correspondence:

Dr William Bernal

Consultant Liver Intensivist

Institute of Liver Studies

Kings College Hospital, Denmark Hill, London SE5 9RS, United Kingdom

Phone: +44 203 299 3367, Fax: +44 203 299 3899

Email: William.bernal@kcl.ac.uk

Running Title:

A diagnostic Model for liver graft non-function

Abbreviations:

AIH: autoimmune hepatitis, ALF: acute liver failure, ALT: alanine aminotransferase, AST: aspartate aminotransferase, AUROC: area under receiver operating curve. ALD: alcohol-related

liver disease, AUROC: area under receiver operating curve. CI: confidence interval, CIT: cold ischaemia time, CTP: Child-Turcotte-Pugh, DBD: donation after brain death, DCD: donation after circulatory death, DRI: donor risk index, EGD: early graft dysfunction. ESLD: end stage liver disease, HBV: hepatitis B virus. HCC: hepatocellular carcinoma, HCV: hepatitis C virus. HR: hazard ratio, AIH: autoimmune hepatitis. HCC: hepatocellular carcinoma. INR: international normalised ratio. LITU: liver intensive therapy unit, LRT: liver retransplantation. LT: liver transplantation, MELD: model for end stage liver disease. NHS: National Health Services, OPTN: Organ Procurement and Transplant Network, PNF: primary non-function. ROC: receiver operating curve. RRT: renal replacement therapy, UK: United Kingdom, UKELD: United Kingdom end-stage liver disease model. UNOS: United Network of Organ Sharing. USA: United States of America, WHO: world health organization, WL: waiting list.

Abstract:

Current diagnostic criteria for primary non-function (PNF) of liver grafts are based on clinical experience rather than statistical methods. A retrospective, single centre study was conducted of all adults (n=1,286) who underwent primary liver transplant 2000-2008 in our centre. Laboratory variable during the first post-liver transplant week were analysed. Forty two patients (3.7%) had 2-week graft failure. Transplant albumin, day-1 AST, day-1 lactate, day-3 bilirubin, day-3 INR and day-7 AST were independently associated with PNF on multivariate logistic regression. PNF score =(0.000280*D1AST)+(0.361*D1 Lactate)+(0.00884*D3 Bilirubin)+(0.940*D 3 INR)+(0.00153*D7 AST)-(0.0972*TxAlbumin)-4.5503. ROC analysis showed the model AUROC of 0.912 (0.889 -0.932) was superior to the current United Kingdom (UK) PNF criteria of 0.669 (0.634-0.704, p<0.0001). When applied to a validation cohort (n=386, 34.4% patients) the model had AUROC of 0.831 (0.789 - 0.867) compared to the UK EGD criteria of 0.674 (0.624-0.721). The new model performed well after exclusion of patients with marginal grafts and when modified to include variables from the first three post-LT days only (AUROC of 0.818, 0.776-0.856, p=0.001). This model is superior to the current UK PNF criteria and is based on statistical methods. The model is also applicable to recipients of all types of grafts (marginal and non-marginal).

Introduction:

Graft injury is inevitable during the process of liver transplantation (LT) to which ischaemia and reperfusion is an important contributor. (1) In the majority of patients, the manifestations of such graft injury settle rapidly within the first week following LT. However, in few patients measures of graft function display significant abnormalities which are referred to as early graft dysfunction (EGD). Patients who develop EGD have increased risk of graft failure or death early in the post-LT period. (2) In its most severe form, primary non-function (PNF), the risk of death or the requirement for emergency liver retransplantation (LRT) within the first 1-2 weeks are markedly increased. (3) (4)

There is a lack of agreement on the terminology used to describe PNF. (5) The diagnostic criteria of PNF also varied significantly in the reported literature. (4) (6) (7) (8) (3) (9) (10) Initial studies suggested the use of markers of liver injury such as aspartate aminotransferase (AST) (3) (4) (6) or alanine aminotransferase (ALT), (3) (6) (7) (8) or markers of hepatic synthetic function such as prothrombin time,(4) (6) (7) (8) bilirubin level, (2) bile production,(6) (8) acidosis, (11) ammonia level (4) or hepatic encephalopathy. (2) The timing of measurement of these variables also varies in the literature with variables included on day-1 post-LT,(3) first 3 days, (8) (12) to those recorded at any day between day 2-7 post-transplant. (4) Furthermore, the time frame of the outcome of PNF varied in the described studies with death or retransplantation occurring within 1 week, (6) day 2-7, (4) within 10 days (11) or up to 2 weeks post-LT. (3) (5)

In the United Kingdom (UK), the National Health Services (NHS) Blood and Transplant Advisory Committee set the criteria for PNF under category 9 of super-urgent listing for liver transplantation. (13) These criteria were based mainly on clinical experience utilising previously reported data (Box 1). In the United States of America (USA), the Organ

Procurement and Transplant Network (OPTN) described a different set of diagnostic criteria for PNF (Box 2). (14) In both these criteria, cut-off points were arbitrary with significant difference in the choice of thresholds for the 2 systems. Moreover, the UK criteria include bile production, based on previous practices of LT. Routine use of biliary drains to monitor bile output is rarely in use nowadays.

The aims of this study were 3 fold: Firstly, to define diagnostic criteria of PNF using standard and transportable laboratory tests performed within the first week post-LT based on statistical models, secondly, to compare the diagnostic performance of the proposed criteria with the current UK criteria for PNF; and lastly, to internally validate the newly proposed criteria.

Box 1: NHS Blood and Transplant Liver Advisory Group diagnostic criteria for super-urgent listing for liver transplantation for early graft dysfunction. **Box 2:** OPTN urgent listing criteria for primary non-function.

Box 1:

Early graft dysfunction on days 0 to 7 after liver transplantation with at least two of the following:

1. AST >10,000 IU/l
2. INR >3.0
3. Serum lactate >3 mmol/l
4. Absence of bile production.

Box 2:

Primary non-function of a transplanted liver within 7 days of implantation; as defined by (a) or (b):

(a) AST ≥ 3,000 and one or both of the following:

- INR ≥ 2.5
- Acidosis, defined as having an arterial pH ≤ 7.30 or venous pH of 7.25 and/or Lactate ≥ 4 mmol/l

(b) Anhepatic candidate

Materials & Methods:

Patients & Design:

We performed a retrospective analysis of all adult patients (n=1,286) who had LT in our centre and received grafts from deceased donors between January 2000 to December 2008. Exclusions included retransplanted patients (n=133), those who died during the transplant operation (n=4), had hepatic artery thrombosis (n=19) as a cause of graft failure and 5 patients who had early death secondary to metastatic cancers. Therefore, data was analysed on 1125 patients.

Dataset:

All patients who had LT in our centre were admitted from operating theatre to liver intensive therapy unit (LITU). All patients admitted to LITU had daily clinical, physiological, laboratory variables and requirement for organ support prospectively recorded into LITU database. This dataset, electronic patient record and the clinical notes were utilised to summate demographic, clinical, and laboratory data at the time of listing, at time of transplant and for the first week following LT. Variables determined daily included AST, bilirubin, INR, lactate, creatinine, requirement for vasopressors and renal replacement therapy (RRT).

Prognostic models were calculated at the time of listing and at transplant. Model for end stage liver disease (MELD) score was calculated according to the United Network of Organ Sharing (UNOS) adjustments. (15) The UK end stage liver disease (UKELD) score was calculated according to Barber et al. (16) Donor and graft variables assessed were age, gender, donor-to-recipient gender mismatch, ethnicity, donor-to-recipient blood group mismatch, height, weight, body mass index (BMI), graft type: donation after brain death (DBD) versus donation after circulatory death (DCD), organ type (whole versus non-whole) and cold ischemia time (CIT). The donor risk index (DRI) was calculated according to Feng et al. (17) Marginal grafts were defined as grafts with DRI > 1.8. (18) Grafts were also categorised as marginal and non-marginal according to surgical inspection of the grafts by the unit's transplant surgeons.

Definitions of outcome measures:

Patient survival was documented according to their recorded survival status in our hospital information system. According to previously published reports, we chose death or the requirement for LRT within 2 weeks of primary liver transplant (PLT) as the primary outcome to define PNF (3) (5) (19) .

Ethical approval:

Ethical approval for interrogation, analysis and publication of this anonymised dataset was obtained from the Southeast London Research Ethics Committee 3 (previously known as King's College Hospital Research Ethics Committee).

Statistical Analysis:

Data were assessed for normality using the D'Agostino Pearson test. Comparisons were made by Student's t-test (one way ANOVA) or Mann Whitney (Kruskall-Wallis) for 2 (or more) comparison groups. Categorical data were compared using the Chi square (χ^2) test. Univariate and multivariate (backwards mode) logistic regression and area under the receiver operating curves (AUROC) analysis were performed for 2-week outcomes. To assess goodness of fit, the Hosmer and Lemeshow test was calculated where a p-value of <0.05 was used to reject variables or models for failing to adequately pass the comparison of observed and expected outcomes due to over-fitting. In multivariate regression, backwards mode, the criterion for exclusion was a p-value of 0.15 to balance the need for a parsimonious model while reducing potential bias from suppressor effects. Therefore, a three-stage filtering process was used where a variable was removed if a) it was radically alterable by therapy, b) if it failed the Hosmer and Lemeshow test ($p<0.05$) and c) if the p value was >0.15 on backward regression. Data were assessed for multiple collinearity using correlation coefficients. For an $R>0.4$ the predictor with the highest regression coefficient was retained. In order to minimise reductions in specificity while attempting to improve sensitivity, specificity was fixed a priori at 95% in deciding cut-offs for any logistic regression models. A second repeated measures logistic regression model was fitted where all data were added to the model to assess for any bias in the above variable selection method as described in the supplementary material. Significance was required at the 95% level with a 2 tailed p-value

of <0.05. Data were analysed in Excel (Microsoft), SPSS 17 (IBM) and MedCalc v 12.2.0 (Mariakerke, Belgium).

RESULTS

Recipient, donor and graft descriptive results:

Forty-two patients (3.7%) had graft failure within 2 weeks of PLT. Twenty-four patients (2.1%) died and 18 patients (1.6%) had LRT within 2 weeks of PLT. Recipient characteristics are summarised in table 1. There were no significant differences in demographic variables or aetiology of liver disease between patients who had 2-week graft failure compared to those who did not, except for a higher proportion of patients transplanted with acute liver failure (ALF) in the 2-week graft failure group. At transplant, patients with graft failure had lower albumin level, higher bilirubin, creatinine, INR, sodium (Na) levels and MELD score. Donor and graft variables were not significantly different between the 2 groups except for higher proportion of ABO mismatch, gender mismatch and higher donor body mass index (table 1). Grafts assessed as non-optimal had a higher DRI than those assessed as optimal 1.68 (1.00-3.70) vs 1.61 (0.93-3.90), p=0.02. Non-optimal livers had higher day-1 AST 663 (39-13886) vs 568 (27-11140), p=0.001. However, no relationship between PNF and graft weight, graft assessment or blood group was observed. On univariate analysis these variables did not reach significance and therefore were not included in subsequent multivariable analysis. DRI, proportion of patients with marginal grafts, DCD grafts, partial grafts and CIT were not significantly different between the two groups.

By expanding the definition of PNF to 2 weeks in total 42 cases were identified compared to 27 who met the USA definition in this time frame and only 14 who met the UK definition in this time frame.

Post-transplant biochemical data:

Lactate, AST, bilirubin, INR and creatinine showed a general trend of decline within the first post-transplant week except for bilirubin and creatinine levels for patients with 2-week graft failure (figure 1). Lactate, AST, bilirubin, INR and creatinine were significantly higher in patients who had 2-week graft failure compared to those whom the graft survived for more than 2 weeks (table 2).

Derivation and validation cohorts

A randomly generated index was used to split the data into derivation and validation cohort (65% and 35% of the entire cohort, respectively). Data were matched ($p>0.05$) between the derivation and validation sets for >95% of variables (see Supplementary table).

Factors associated with PNF:

Univariate and multivariate logistic regression analysis are shown in Table 3. There were a large number of variables which on univariate analysis could be initially justified in progressing to multivariate analysis based on p-value alone. Of note, DRI and DCD status did not have univariate significance for predicting this short-term outcome. We employed two methods to help simplify the regression model. First, if a variable was prone to bias in terms of being potentially modifiable by treatment (eg. serum creatinine or decision to continue/start RRT) and second, statistically, if the Hosmer and Lemshow p-value was <0.05 there was a danger of poor fit in any subsequent model. Using these filters, we identified transplant albumin, day-1 AST, day-1 lactate, day-3 bilirubin, day-3 INR and day-7 AST as independent factors associated with PNF. Day-1 and day-7 AST were not closely correlated ($r=0.05$, $p=0.010$). MELD score at transplantation as a continuous or categorical variable with cut off value of 25 (20) (21)or 30 (22) was not associated with PNF on multivariate

analysis. Based on these findings, we constructed a model composed of the aforementioned variables in the first week post-LT. This PNF score =(0.000280*D1AST)+(0.361*D1 Lactate)+(0.00884*D3 Bilirubin)+(0.940*D 3 INR)+(0.00153*D7 AST)-(0.0972*TxAlbumin)-4.5503.

Diagnostic performance of the PNF model compared to existing models:

We compared our model with the UK category 9 criteria for super urgent listing (box 1) and with the USA OPTN diagnostic criteria for PNF (box 2). Figure 2 illustrates ROC analysis that demonstrates superior performance of the new PNF diagnostic model with AUROC of 0.912 (0.889 -0.932) compared to the UK EGD criteria of 0.669 (0.634-0.704, $p<0.0001$ for pairwise comparison to the new model) and the USA PNF criteria of 0.776 (0.774-0.806, $p=0.010$ for pairwise comparison to the new model). Our model also showed significant improvement in the sensitivity of the model in detecting patients with PNF (73%) compared to the UK (31%) and USA (66%) criteria (table 4).

Validation of the new model:

The same model was applied to the validation cohort of 386 (34.4%) randomly selected patients. The proposed model had AUROC of 0.831 (0.789 -0.867) compared to the UK EGD criteria of 0.674 (0.624-0.721) and the USA PNF criteria of 0.781 (0.736-0.822) as shown in figure 2.

Earlier detection of PNF:

Giving the importance of earlier detection and diagnosis of PNF to be able to offer patients with PNF a curative treatment (retransplantation), we modified the above model to include variables from the first 3 days post-transplant only. This PNF score =(0.000280*D1 AST)+(0.361*D1 Lactate)+(0.00884*D3 Bilirubin)+(0.940*D3 INR)-(0.0972*TxAlbumin)-4.5503. This model also performed well with AUROC of 0.818 (0.776-0.856, $p=0.001$).

Pairwise comparison of AUROC of the two models (including and excluding day-7 variables) showed small and non-significant change of 0.83 v 0.82, p=0.321, figure 4).

Application of the model to recipients of marginal grafts:

Published reports on this topic varied with regards to the inclusion or exclusion of groups of patients such as patients with ALF, recipients of DCD grafts and recipients of partial or split grafts. We compared the diagnostic performance of our model on the entire validation cohort and after exclusion of each of the sub groups mentioned above as illustrated in figure 3. The new model continued to perform well after exclusion of recipients of DCD grafts and recipients of partial grafts but there was a drop in the AUROC from 0.840 to 0.795 after exclusion of patients transplanted for ALF. Despite the decrease in AUROC, our model continued to discriminate well between those who met the outcome definition of PNF from those who did not (p=0.041).

Effect of renal support

When modelled with the dichotomous variable of whether the patient received haemo-filtration on that day, only CVVHF on day 7 was independently associated with PNF and therefore added into the model. The modified model had an AUROC of 0.828 so there was no benefit in its addition.

Addition of serum creatinine from all measured post-transplant days did not improve the model and none of the creatinine measurements were statistically significant in terms of prognostic performance when forced into logistic regression models

Assessment of proposed model versus repeated measures logistic regression

We also performed a repeated measures logistic regression model, initially with all IV entered. The final model included AST, lactate, bilirubin and requirement for vasopressor therapy with highly significant predictive accuracy and a sensitivity of 71 (55-84)% and specificity of 95 (92-97)%. This was essentially the same as with the original proposed model.

DISCUSSION:

The diagnostic criteria for primary non-function in the UK are described under category-9 of super-urgent listing for retransplantation under the title of early graft dysfunction. (13) These criteria were based on clinical experience and liver transplant practices of approximately 2 decades ago. During the last 20 years, significant progress was achieved in the field of liver transplantation such as improvements in operative methods, immunosuppression and post-operative ICU care. (23) This resulted in the current 1- and 5-year survival rates post-transplant of 90% and > 50%, respectively. Despite this improvement, 2.7-6.9% can develop primary non-function with high early post-transplant mortality without retransplantation. (3) (4) (7)

This is the largest study to date to investigate the diagnostic criteria of primary liver graft non-function. This study has a number of design strengths compared to previous reports which are worth exploring. (3) (4) (7) (8) First, we set our diagnostic method based on composite of outcome variables of death and retransplantation rather than categorising graft function according to laboratory variables. Second, we chose the time interval of the composite outcome measure of 2-week post-transplant to include patients who met the

current UK diagnostic criteria by the end of the first week who would have been retransplanted or died within the second week. This was also in agreement with Strasberg et al. suggestion on extending the time interval for the diagnosis of PNF from 1-week to 2-week outcome in an attempt to refine the definition of PNF; and in agreement with other authors. (3) (5) (19) Third, this study focused on adult patients to avoid any heterogeneity in the patient population whilst previous studies included adults and paediatric patients. Fourth, we excluded patients who developed graft dysfunction or died because of HAT, those who died during the transplant procedure and patients who developed graft dysfunction or died following a second or subsequent transplant which homogenise the patient population.

The rate of PNF according to this paper definition was 3.7% which is significantly lower than 5.8% reported by Johnson and colleagues utilising SRTR database of > 10,000 patients and lower than 6-9.2% reported by single centre series. (19) (24) (25) We identified a large number of variables which were associated with graft failure on univariable analysis in addition to those documented in table 4 such as requirement for RRT, creatinine level, requirement for vasopressors and mechanical ventilation during the first week post transplant. However, we excluded modifiable variables such as the requirements for RRT, creatinine level or vasopressor support and other variables which might simply reflect a severity of illness such as requirement for mechanical ventilation in critically ill patients . Other variables were excluded according to statistical methods (goodness of fit).

There is evidence that the severity of recipient illness may have an impact on early post-transplant outcomes. High MELD score at transplantation (cut off values 25 or30) was associated with reduced 3- and 12-month post-transplant survival. (20) (21) (22) (26) Our analysis showed that neither MELD (as continuous or categorical variable) nor its components were associated with PNF after controlling for other pre- and post-transplant variables. This finding was consistent with Johnson et al. who used the same definition of

PNF in an analysis of SRTR database of 10,545 patients and showed no association of MELD with PNF. Furthermore, the rate of PNF after the implementation MELD score for listing in USA was relatively stable at approximately 6% which indicates that despite transplanting sicker patients, the rate of PNF did not change. (19)

However, our model included serum albumin at transplantation as a surrogate marker of recipient severity of illness. Serum albumin is one of the five components of the well established Child-Turcotte-Pugh model, a system used over four decades to assess severity of illness in patients with chronic liver disease. (27) More recently, serum albumin was found to have inverse linear relationship to waiting list mortality analysing UNOS database and in a single Canadian transplant cohort. Albumin was independent of MELD in predicting waiting list mortality. Moreover, the addition of serum albumin to MELD and MELDNa (5 variable MELD) significantly improved its ability to identify patients at risk of death on the transplant waiting list. (28) (29) (30).

We constructed our model for the diagnosis of PNF based on easily obtainable laboratory parameters during the first post-transplant week. Our multivariable regression model identified albumin level at transplant, day-1 blood lactate and AST levels, day-3 bilirubin and INR levels, and day-7 AST as factors independently associated with PNF. Blood lactate is already in use as part of the diagnostic criteria for PNF in the UK and USA. (13) (14) Blood lactate is a bi-product of anaerobic metabolism secondary to abnormal tissue microcirculatory perfusion and oxygenation. (31) It is mainly metabolised by the liver and therefore, higher blood lactate levels may reflect both increased production and impaired clearance in patients with hepatic dysfunction. (32) Unsurprisingly, AST and INR were included in our model based on statistical methods. High AST levels post-transplant indicates acute graft injury and INR reflects the graft synthetic function and are both included in the current UK and USA criteria for PNF. (13) (14) We identified bilirubin level as an

independent factor associated with PNF. Although bilirubin is not included in the current UK or USA criteria for PNF, bilirubin was associated with graft dysfunction in previous reports. (2) (9) (33)

It is believed that earlier liver retransplantation of patients with graft failure within the first post-transplant week carries higher survival compared to retransplantation between the second to fourth weeks. (34) (24) Giving the importance of early diagnosis of PNF, we modified our model to include variables from the first 3 post-LT days. The modified model showed good diagnostic performance with AUROC of 0.82 which was not significantly different to AUROC of 0.83 for the original model which utilises day-7 variables. This indicates that the modified model can detect patients with PNF as early as the third post-operative day. Should these criteria become utilised for super-urgent listing, it can facilitate earlier listing of patients with PNF for LRT.

The diagnostic performance of our model was excellent when applied to the derivation cohort with AUROC 0.912. Although there was a reduction in the diagnostic performance of the model when applied to the validation cohort with AUROC of 0.83, the model had very good diagnostic performance compared to exiting UK and USA diagnostic criteria. (35) We demonstrated a significant improvement in the sensitivity of this model (73%) compared to the existing UK (31%) and USA diagnostic criteria for PNF (66%). This was achieved without significant reduction in specificity of 95% compared to specificity of 93% and 98% for both USA and UK criteria, respectively. This model also has significantly improved negative likelihood ratio (LK-) of 0.3 compared to existing models of 0.7 and 0.6 for UK and USA PNF models, respectively indicating superior discrimination of our model in ruling out cases who did not meet the outcome definition of 2-week graft failure, assuming the model is subsequently validated. (36)

Initial poor function of the graft may be attributed to donor or graft variables. (3, 4, 37) Therefore, we applied our model to the validation cohort before and after exclusion of recipients of DCD organs and patients who received partial grafts. The performance of the new model assessed by ROC characteristics did not alter significantly with AUROC of 0.840, 0.854 and 0.821, respectively. This indicates that this model can be applied to recipients of all graft types (marginal and non-marginal) without significant difference in its diagnostic performance. Furthermore, none of the donor or graft variables including the use of marginal grafts with DRI > 1.8 was associated with PNF on logistic regression. The reported literature varied with regards to the impact of graft and donor quality on the development of PNF. Makowka et al. found no impact of donor variables on the development of PNF. (7) Others found split or reduced grafts, advanced fatty changes in the graft, longer cold or warm ischaemia times, older donor age, donor renal insufficiency, ICU length of stay > 3 days, donor weight > 100kg, duration of the anhepatic phase and DCD grafts were associated with poor graft function. (3) (4) (19, 37) (38) (39) Findings of these reports are difficult to compare to ours for a number of reason. Firstly, different local practices of donor recipient matching, secondly, different definitions of PNF (as diagnostic criteria or time frame) and lastly different era of transplant practices in some of these studies.

There are several limitations to this study. Firstly, it represents a single centre experience and therefore, applicability of the results to other patient populations may be limited. Secondly, operative data such as operation time, volume of blood loss, intra-operative transfusion requirements, preservation solutions used and warm ischaemia time were not available to us and accordingly were not included in our analysis. These variables could be associated with early post-transplant outcomes. (3) (4) (8) (40) Finally, data on immunosuppression were not included in this study; however, all patients received standard immunosuppression according to our institution guidelines.

It is interesting that using a more formal approach to repeated measures without significant supervised variable selection led to a model with similar prognostic accuracy. The model we propose is a relatively simple equation and unless these more data driven techniques can provide much higher accuracy we do not recommend them at present. Also, they require data to day 7 and our model can allow stratification at day 3 where prompt preparations for re-transplantation can be started.

While we performed several steps to statistically validate the findings of our analysis and model this should not be taken as a final external validation. Particularly since we define PNF in a novel manner then an assessment of improved accuracy and/or survival at several centres would need to be made before changes in practice can be recommended. We make a distinction between PNF and end stage graft dysfunction beyond 2 weeks' post-transplant which is likely to present as a different clinical entity (putatively called early graft dysfunction). For our PNF model rapid assessment and decisions regarding re-transplant can be biased by clinical decisions made outside the modelled parameters and therefore we invite and recommend more robust external validation of the increased sensitivity suggested by our model.

Given that there are differing definitions of PNF between Europe, North American and Asia it would be particularly useful to validate the findings of this score in other centres firstly in the UK and then internationally to assess any bias both from our approach and those where relisting decisions for PNF use alternative criteria.

In conclusion, this study addresses a life-threatening complication of liver transplantation (PNF) and revisits the diagnostic criteria of this complication in the MELD and UKELD era based on acceptable outcome definition. Our proposed criteria utilise easily obtainable and

objective laboratory parameters and it is based on statistical methods. Further, we have shown that our criteria can be modified to obtain earlier diagnosis of PNF with preservation of its diagnostic performance. Our diagnostic criteria are also applicable to recipients of all types of grafts (marginal and non-marginal). We invite the transplant community in the UK to externally validate our model to improve the current diagnostic criteria for PNF.

Table 1: Recipient, graft and donor characteristics for all patients and comparison of variables between patients who met the definition of 2-week graft failure and those who did not.

Recipient parameters	All patients n=1125	2-week graft failure		p-value
		No, n=1083	Yes, n=42	
Demographic				
Age (years)	51 (16-74)	51(16-74)	48(18-68)	0.196
Gender: male (%)	679 (60.4)	657 (60.7)	22 (52.4)	0.282
Hepatocellular carcinoma (%)	177 (15.7)	173 (16.0)	4 (9.5)	0.260
Aetiology				
Viral (%)	286 (25.4)	276 (25.5)	10 (23.8)	0.807
ALD (%)	225 (20.0)	217 (20.0)	8 (19.0)	0.875
AIH / Cholestatic (%)	215 (19.1)	212 (19.6)	3 (7.1)	0.045
Acute liver failure	177 (15.7)	161 (14.9)	16 (38.1)	<0.001
Cryptogenic (%)	71 (6.3)	69 (6.4)	2 (4.8)	1.000
Other (%)	151 (13.4)	148 (13.7)	3 (7.1)	0.353
Transplant laboratory variables				
Albumin (g/dl)	2.9 (0.7-5.6)	2,9 (0.7-5.6)	2.4 (0.9-4.4)	<0.001
Bilirubin mg/dl	3.0 (0.1-53.9)	3.0 (0.1-53.9)	3.7 (0.5-31.1)	0.044
Creatinine µmol/l	1.1 (0.4-11.5)	1.1 (0.4-11.5)	1.2 (0.8-4.8)	0.004
INR	1.34 (0.83-16.00)	1.3 (0.8-16.0)	1.7 (1.0-15.0)	0.002
Sodium (meq/l)	137 (112-159)	137 (112-159)	139 (126-150)	0.014
MELD	16 (6-40)	16 (6-40)	21 (7-40)	0.002
UKELD	55 (41-80)	55(41-77)	56(49-80)	0.060
Donor & graft variables				
Age (years)	47 (11-82)	47 (11-82)	45 (16-76)	0.532
Gender: male (%)	564 (50.1)	542 (52.0)	22 (53.7)	0.831
Ethnicity: European (%)	1022 (96.1)	984 (96.2)	38 (92.7)	0.217
Gender mismatch (%)	456 (42.1)	432 (41.4)	24 (58.5)	0.029
ABO blood group mismatch (%)	94 (8.5)	84 (7.9)	10 (23.8)	<0.001
Donor height (cm)	170 (130-198)	170 (130-198)	170 (148-193)	0.645
Donor weight (kg)	75 (30-140)	75 (30-140)	75 (54-117)	0.076
BMI (kg/m^2)	25 (12-47)	25 (12-47)	26 (18-35)	0.030
Graft weight (g)	1455 (230-2815)	1480 (916-2220)	1461 (230-2815)	
Donor cause of death: CVA (%)	738 (65.6)	711 (66.9)	27 (64.3)	0.726
Graft type: split or reduced (%)	143 (12.9)	136 (12.8)	7 (16.7)	0.460
Donation after cardiac death (%)	91 (8.2)	87 (8.2)	4 (9.5)	0.772
Cold ischaemia time (hours)	10.1 (0.9-20.5)	10.1 (0.9-20.5)	10.6 (5.2-19.1)	0.334
DRI	1.7 (0.9-3.7)	1.7 (0.9-3.7)	1.7 (1.1-2.7)	0.568
Non-optimal grafts	458 (41.6)	441 (41.6)	17 (41.5)	0.889
Marginal grafts-DRI > 1.8 (%)	297 (26.4)	286 (26.4)	11 (26.2)	0.972

Table 2: Comparison of post-transplant biochemical profile according to 2-week graft failure.

Variable	2-week graft failure		p-value
	No, n=1083	Yes, n=42	
D1 AST (IU/l)	612 (27-13,886)	1650 (34-12,200)	<0.001
D1 Bilirubin (mg/dl)	2.9 (0.1-38.5)	4.3 (0.5-33.2)	0.010
D1 Creatinine (mg/dl)	1.2 (0.1-6.3)	1.5 (0.8-3.7)	<0.001
D1 INR	1.4 (0.9-7.6)	1.7 (1.0-9.8)	<0.001
D1 Lactate	1.3 (0.3-11.2)	3.2 (0.9-17.0)	<0.001
D2 AST (IU/l)	301 (21-8288)	1209 (47-8078)	<0.001
D2 Bilirubin (mg/dl)	2.3 (0.1-32,3)	5.1 (0.6-33.2)	<0.001
D2 Creatinine (mg/dl)	1.3 (0.4-6.8)	1.7 (0.8-3.2)	0.002
D2 INR	1.2 (0.5-5.1)	1.4 (0.9-6.9)	<0.001
D2 Lactate (mmol/l)	1.2 (0.3-10)	2.0 (0.6-20)	<0.001
D3 AST (IU/l)	154 (1-4223)	774 (49-5365)	<0.001
D3 Bilirubin (mg/dl)	2.7 (0.1-22.4)	5.4 (0.4-25.1)	<0.001
D3 Creatinine (mg/dl)	1.2 (0.4-20.9)	1.6 (0.7-4.1)	<0.001
D3 INR	1.1 (0.8-4.1)	1.4 (0.9-4.4)	<0.001
D3 Lactate (mmol/l)	1.1 (0.4-10.0)	2.0 (0.7-10.5)	<0.001
D4 AST (IU/l)	92 (12-3408)	395 (20-2328)	<0.001
D4 Bilirubin (mg/dl)	3.3 (0.1-37.2)	6.6 (0.6-21.7)	<0.001
D4 Creatinine (mg/dl)	1.1 (0.2-8.0)	1.7 (0.2-3.4)	<0.001
D4 INR	1.1 (0.8-4.4)	1.2 (0.9-7.0)	<0.001
D4 Lactate (mmol/l)	1.2 (0.3-11.2)	1.9 (0.8-12.7)	<0.001
D7 AST (IU/l)	62 (9-3662)	158 (22-10040)	<0.001
D7 Bilirubin (mg/dl)	2.9 (3-760)	6.3 (0.8-21.9)	<0.001
D7 Creatinine (mg/dl)	1.1 (0.5-7.4)	1.6 (0.7-3.4)	<0.001
D7 INR	1.0 (0.8-3.1)	1.1 (0.9-3.9)	<0.001
D7 Lactate (mmol/l)	1.4 (0.4-8.8)	2.2 (0.8-24.9)	<0.001

Table 3: Factors associated with 2-week graft failure on univariate and multivariate logistic regression statistics in the randomly selected derivation cohort (n=739, 65.7%).

	Univariate			Multivariate		
Variables	**OR**	**95% CI**	**P-value**	**OR**	**95% CI**	**P-value**
Demographic						
Age	0.973	0.948-0.999	0.039			
Gender	1.207	0.568-2.567	0.625			
Aetiology						
ALD	1.043	0.417-2.611	0.928			
Viral	0.937	0.394-2.230	0.883			
AIH/Cholestatic	0.307	0.072-1.307	0.110			
ALF	3.728	1.710-8.127	0.001			
HCC	0.364	0.085-1.552	0.172			
At transplant						
Albumin (g/dl)	0.914	0.870-0.960	<0.001	0.9073	0.8429-0.9766	0.001
Bilirubin (mg/dl)	1.002	1.000-1.004	0.030			
Creatinine (mg/dl)	1.002	0.999-1.005	0.113			
INR	1.15	1.033-1.279	0.011			
Na	1.058	0.997-1.123	0.065			
MELD	1.058	1.022-1.095	0.001			
MELD >25	2.369	1.108-5.062	0.026			
MELD >30	3.892	1.808-8.378	0.001			
UKELD	1.052	0.998-1.109	0.061			
Graft and donor						
Age	0.989	0.965-1.013	0.36			
ABO mismatch	4.855	2.049-11.505	<0.001			
Donor weight (kg)	1.033	1.007-1.061	0.014			
BMI (kg/m^2)	1.099	1.020-1.185	0.013			
Graft weight	1.005	0.999-1.001	0.332			
CIT	1.083	0.957-1.226	0.208			
Split or reduced grafts (%)	1.012	0.345-2.971	0.983			
DCD grafts	1.308	0.384-4.453	0.668			
Optimal vs non-optimal	0.992	0.572-1.869	1.00			
DRI	1.006	0.369-2.743	0.991			
DRI >1.8	1.001	0.436-2.297	0.998			
Post transplant						
D1 AST (IU/l)	1.001	1.000-1.001	<0.001	1.0003	1.0000-1.0005	0.039
D1 Bilirubin (mg/dl)	1.006	1.002-1.009	<0.001			
D1 Creatinine (mg/dl)	1.005	1.001-1.010	0.012			
D1 INR	2.338	1.639-3.335	<0.001			
D1 Lactate (mmol/l)	1.627	1.397-1.896	<0.001	1.4351	1.1813-1.7435	<0.001
D1 Na (meq/l)	1.093	1.008-1.186	0.032			
D2 AST (IU/l)	1.001	1.000-1.001	<0.001			
D2 Bilirubin (mg/dl)	1.008	1.004-1.011	<0.001			

D2 Creatinine (mg/dl)	1.003	0.999-1.008	0.117			
D2 INR	2.548	1.622-4.002	<0.001			
D2 Lactate (mmol/l)	1.881	1.442-2.454	<0.001			
D2 Na (meq/l)	1.075	1.002-1.153	0.044			
D3 AST (IU/l)	1.001	1.001-1.002	<0.001			
D3 Bilirubin (mg/dl)	1.01	1.005-1.014	<0.001	1.0089	1.0031-1.0147	0.003
D3 Creatinine (mg/dl)	1.001	0.999-1.004	0.290			
D3 INR	4.807	2.514-9.189	<0.001	2.5611	1.1855-5.5327	0.017
D3 Lactate (mmol/l)	1.578	1.287-1.936	<0.001			
D3 Na (meq/l)	1.118	1.047-1.193	0.001			
D4 AST (IU/l)	1.003	1.002-1.004	<0.001			
D4 Bilirubin (mg/dl)	1.006	1.002-1.010	0.002			
D4 Creatinine (mg/dl)	1.003	1.000-1.007	0.061			
D4 INR	22.179	6.455-76.206	<0.001			
D4 Lactate (mmol/l)	1.781	1.295-2.450	<0.001			
D4 Na (meq/l)	1.156	1.081-1.236	<0.001			
D7 AST (IU/l)	1.002	1.001-1.003	0.005	1.0015	1.0005-1.0025	0.003
D7 Bilirubin (mg/dl)	1.004	1.000-1.007	0.029			
D7 Creatinine (mg/dl)	1.005	1.001-1.008	0.008			
D7 INR	16.353	5.294-50.517	<0.001			
D7 Lactate (mmol/l)	1.475	1.119-1.944	0.006			
D7 Na (meq/l)	1.107	1.027-1.193	0.008			

Table 4: Comparison of the performance of the proposed model to the current UK and USA models.

Factor	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	Accuracy
Proposed model	73 (39-94)%	95 (92-97)%	13.4 (7.6-23.3)	0.3 (0.1-0.8)	0.909
UK Category 9 criteria (PNF)	31 (9-61)%	98 (97-99)%	19 (6-60)	0.7 (0.5-1)	0.971
USA OPTN PNF criteria	66 (46-82)%	93 (92-95)%	10.6 (7.2-15.6)	0.4 (0.2-0.6)	0.926

Supplementary table: Comparison of derivation and validation cohorts.

Variables	Derivation	Validation	P-value
Total = 1109	727	382	
Liver related death or LRT within 2 weeks (Y:N)	23:704	10:372	0.747
Demographic			
Age	51(16-72)	51(17-74)	0.990
Gender (male:female)	457:270	213:169	0.025
Aetiology			
Viral (Y/N))	183:544	99:283	0.583
ALF (Y:N)	105:622	65:317	0.297
At listing			
Albumin (g/dl)	3,0(0.9-5.0)	3.0(0.9-7.9)	0.868
Bilirubin (mg/dl)	2.6(0.1-43.4)	2.6(40.2-49.7)	0.930
Creatinine (mg/dl)	1.0(0.4-10.1)	1.0(0.5-9.3)	0.622
INR	1.22(0.85-5.82)	1.24(0.77-76.0)	0.348
Na (meq/l)	136(117-149)	137(121-147)	0.504
MELD	14(3-43)	14(2-35)	0.820
UKELD	54(38-75)	55(41-71)	0.897
At transplant			
Albumin (g/dl)	2.9(0.9-5.0)	2.9(0.7-7.1)	0.551
Bilirubin (mg/dl)	2.6(0.1-42.5)	2.6(0.2-49.7)	0.930
Creatinine (mg/dl)	1.1(0.4-11.5)	1.1(0.4-7.8)	0.953
INR	1.31(0.83-16.0)	1.36(0.85-15.0)	0.172
Na (meq/l)	137(112-159)	138(120-157)	0.144
MELD	16(6-59)	17(6-58)	0.251
UKELD	55(41-77)	55(42-73)	0.950
Graft and donor			
Age	47(11-82)	46(11-79)	0.417
BMI (kg/m^2)	24.9(12.3-47.3)	24.7(14.9-47.3)	0.940
CIT (minutes)	604(55-1231)	598(237-1202)	0.800
Split or reduced grafts (Y:N)	99:621	42:330	0.292
DCD grafts (Y:N)	60:660	31:341	0.908
DRI	1.7(0.9-3.7)	1.6(1.0-3.4)	0.071
AST(iU/L)	36(1-786)	38(8-815)	0.122
Post transplant			
D1 AST (IU/l)	625(28-11204)	616(27-13886)	0.916
D1 Bilirubin (mg/dl)	3.0(0.1-38.4)	2.7(0.1-23.8)	0.243
D1 Creatinine (mg/dl)	1.2(0.2-5.8)	1.2(0.1-6.3)	0.337
D1 INR	1.43(0.93-7.60)	1.41(0.92-9.82)	0.727
D1 Lactate (mmol/l)	1.3(0.1-17.0)	1.3(0.1-14)	0.275
D1 Na (meq/l)	143(124-157)	143(128-159)	0.707
D2 AST (IU/l)	307(21-8288)	299(33-7117)	0.491
D2 INR	1.20(0.50-6.9)	1.17(0.80-4.80)	0.370
D2 Lactate (mmol/l)	1.2(0.4-20)	1.20(0.3-10)	0.500

D2 Na (meq/l)	140(124-163)	141(95-157)	0.217
D3 AST (IU/l)	164(1-5635)	144(27-4223)	0.223
D3 Bilirubin (mg/dl)	2.7(0.1-25.0)	2.6(0.3-19.9)	0.289
D3 Creatinine (mg/dl)	1.2(0.5-20.9)	1.2(0.5-7.2)	0.364
D3 INR	1.1(0.85-4.41)	1.1(0.9-3.2)	0.189
D3 Lactate (mmol/l)	1.1(0.1-10.5)	1.1(0.1-7.9)	0.965
D3 Na (meq/l)	138(110-157)	139(115-160)	0.322
D4 AST (IU/l)	95(12-3408)	94(19-2963)	0.506
D4 Bilirubin (mg/dl)	3.4(0.1-37.2)	3.2(0.2-23.6)	0.346
D4 Creatinine (mg/dl)	1.1(0.2-8.0)	1.2(0.5-6.5)	0.555
D4 INR	1.1(0.9-4.4)	1.1(0.8-3.4)	0.472
D4 Lactate (mmol/l)	1.2(0.1-12.7)	1.2(0.3-7.4)	0.757
D4 Na (meq/l)	136(122-157)	137(123-153)	0.127
D7 AST (IU/l)	62(9-10040)	63(12-2098)	0.889
D7 Bilirubin (mg/dl)	3.1(0.2-44.4)	2.8(0.2-36.3)	0.440
D7 Creatinine (mg/dl)	1.1(0.5-7.4)	1.1(0.5-5.0)	0.248
D7 INR	1.1(0.8-3.2)	1.10.8-3.0)	0.742
D7 Lactate (mmol/l)	1.4(0.1-25)	1.3(0.1-17)	0.220
D7 Na (meq/l)	136(120-154)	136(117-154)	0.016

Figure legends:

Figure 1: Sequential changes in main discriminatory biochemical indices over the first week post liver transplantation. Median (95%CI median) values are plotted. A) Lactate B) AST C) Bilirubin D) Creatinine E) INR

Figure 2: A) Derivation cohort comparison of AUROC curves. King's PNF AUROC 0.912 (0.889 - 0.932), USA PNF AUROC 0.776(0.774-0.806, p=0.010 compared to King's PNF using Hanley & McNeil method), UK EGD AUROC 0.669 (0.634-0.704, p<0.001compared to King's PNF using Hanley & McNeil method)

B) Validation cohort comparison of AUROC curves Kings PNF AUROC 0.831 (0.789 -0.867) USA PNF 0.781(0.736-0.822, p=0.547 compared to King's PNF using Hanley & McNeil method) UK EGD 0.674 (0.624-0.721, p=0.154 compared to King's PNF using Hanley & McNeil method).

Figure 3: Comparison of the performance of King's model for diagnosis of primary non-function

a. All patients, AUROC=0.840 (0.799-0.876), p<0.001.

b. Excluding patients with ALF, AUROC=0.795 (0.746-0.839), p=0.041

c. Excluding recipients of DCD grafts, AUROC=0.854 (0.812-0.890, p<0.001

d. Excluding recipients of partial grafts, AUROC=0.821 (0.775-0.861),p=0.003

e. Excluding recipients of marginal grafts, AUROC=0.860 (0.814-0.898),p<0.001

Figure 4: Pairwise comparison of AUROC for King's model of PNF inclusive and exclusive of day-7 variables (AUROC 0.831 v 0.818, p=0.321 for comparison using Hanley & McNeil method).

References:

1. Busuttil RW, Tanaka K. The utility of marginal donors in liver transplantation. Liver Transpl 2003;9(7):651-663.
2. Deschenes M, Belle SH, Krom RA, Zetterman RK, Lake JR. Early allograft dysfunction after liver transplantation: a definition and predictors of outcome. National Institute of Diabetes and Digestive and Kidney Diseases Liver Transplantation Database. Transplantation 1998;66(3):302-310.
3. Mor E, Klintmalm GB, Gonwa TA, Solomon H, Holman MJ, Gibbs JF et al. The use of marginal donors for liver transplantation. A retrospective study of 365 liver donors. Transplantation 1992;53(2):383-386.
4. Ploeg RJ, D'Alessandro AM, Knechtle SJ, Stegall MD, Pirsch JD, Hoffmann RM et al. Risk factors for primary dysfunction after liver transplantation--a multivariate analysis. Transplantation 1993;55(4):807-813.
5. Strasberg SM, Howard TK, Molmenti EP, Hertl M. Selecting the donor liver: risk factors for poor function after orthotopic liver transplantation. Hepatology 1994;20(4 Pt 1):829-838.
6. Carraro P, Varagnolo MC, Cillo U, Tedeschi U, Burra P, Plebani M. Laboratory test scores to aid identification of primary nonfunction of liver transplants. Clin Chem 1995;41(3):471.
7. Makowka L, Gordon RD, Todo S, Ohkohchi N, Marsh JW, Tzakis AG et al. Analysis of donor criteria for the prediction of outcome in clinical liver transplantation. Transplant Proc 1987;19(1 Pt 3):2378-2382.
8. Gonzalez FX, Rimola A, Grande L, Antolin M, Garcia-Valdecasas JC, Fuster J et al. Predictive factors of early postoperative graft function in human liver transplantation. Hepatology 1994;20(3):565-573.
9. Ben-Ari Z, Weiss-Schmilovitz H, Sulkes J, Brown M, Bar-Nathan N, Shaharabani E et al. Serum cholestasis markers as predictors of early outcome after liver transplantation. Clin Transplant 2004;18(2):130-136.
10. Ardite E, Ramos C, Rimola A, Grande L, Fernandez-Checa JC. Hepatocellular oxidative stress and initial graft injury in human liver transplantation. J Hepatol 1999;31(5):921-927.
11. Briceno J, Ciria R. Early graft dysfunction after liver transplantation. Transplant Proc 2010;42(2):631-633.
12. Nanashima A, Pillay P, Verran DJ, Painter D, Nakasuji M, Crawford M et al. Analysis of initial poor graft function after orthotopic liver transplantation: experience of an australian single liver transplantation center. Transplant Proc 2002;34(4):1231-1235.
13. NHS Od. NHS Blood and Transplant Liver Advisory Group. Protocols and guidelines for adults undergoing deceased donor liver transplantation in the UK. 4.1.1 Super-urgent liver transplantation. 2012; NHS Blood and Transplant Liver Advisory Group. Protocols and guidelines for adults undergoing deceased donor liver transplantation in the UK. 4.1.1 Super-urgent liver transplantation. Available from: http://www.organdonation.nhs.uk/ukt/about_transplants/organ_allocation/liver/liver_organ_sharing_principles/liver_organ_sharing_principles.asp#b1. Accessed 24/09/2012]. Available from:
14. OPTN. [cited 2011 02/02/2011]; Organ distribution: allocation of livers. Available from: http://optntransplanthrsagov/policiesandbylaws2/policies/pdfs/policy_8pdf. [Accessed 02/02/2011]]. Available from: http://optntransplanthrsagov/policiesandbylaws2/policies/pdfs/policy_8pdf
15. Wiesner R, Edwards E, Freeman R, Harper A, Kim R, Kamath P et al. Model for end-stage liver disease (MELD) and allocation of donor livers. Gastroenterology 2003;124(1):91-96.
16. Barber K, Madden S, Allen J, Collett D, Neuberger J, Gimson A. Elective liver transplant list mortality: development of a United Kingdom end-stage liver disease score. Transplantation 2011;92(4):469-476.

17. Feng S, Goodrich NP, Bragg-Gresham JL, Dykstra DM, Punch JD, DebRoy MA et al. Characteristics associated with liver graft failure: the concept of a donor risk index. Am J Transplant 2006;6(4):783-790.
18. Thuluvath PJ, Guidinger MK, Fung JJ, Johnson LB, Rayhill SC, Pelletier SJ. Liver transplantation in the United States, 1999-2008. Am J Transplant 2010;10(4 Pt 2):1003-1019.
19. Johnson SR, Alexopoulos S, Curry M, Hanto DW. Primary nonfunction (PNF) in the MELD Era: An SRTR database analysis. Am J Transplant 2007;7(4):1003-1009.
20. Saab S, Wang V, Ibrahim AB, Durazo F, Han S, Farmer DG et al. MELD score predicts 1-year patient survival post-orthotopic liver transplantation. Liver Transpl 2003;9(5):473-476.
21. Desai NM, Mange KC, Crawford MD, Abt PL, Frank AM, Markmann JW et al. Predicting outcome after liver transplantation: utility of the model for end-stage liver disease and a newly derived discrimination function. Transplantation 2004;77(1):99-106.
22. Yoo HY, Thuluvath PJ. Short-term postliver transplant survival after the introduction of MELD scores for organ allocation in the United States. Liver Int 2005;25(3):536-541.
23. Neuberger J, Gimson A, Davies M, Akyol M, O'Grady J, Burroughs A et al. Selection of patients for liver transplantation and allocation of donated livers in the UK. Gut 2008;57(2):252-257.
24. Busuttil RW, Farmer DG, Yersiz H, Hiatt JR, McDiarmid SV, Goldstein LI et al. Analysis of long-term outcomes of 3200 liver transplantations over two decades: a single-center experience. Ann Surg 2005;241(6):905-916; discussion 916-908.
25. Jain A, Reyes J, Kashyap R, Dodson SF, Demetris AJ, Ruppert K et al. Long-term survival after liver transplantation in 4,000 consecutive patients at a single center. Ann Surg 2000;232(4):490-500.
26. Freeman RB, Harper A, Edwards EB. Excellent liver transplant survival rates under the MELD/PELD system. Transplant Proc 2005;37(2):585-588.
27. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg 1973;60(8):646-649.
28. Myers RP, Shaheen AA, Aspinall AI, Quinn RR, Burak KW. Gender, renal function, and outcomes on the liver transplant waiting list: assessment of revised MELD including estimated glomerular filtration rate. J Hepatol 2011;54(3):462-470.
29. Myers RP, Shaheen AA, Faris P, Aspinall AI, Burak KW. Revision of MELD to include serum albumin improves prediction of mortality on the liver transplant waiting list. PLoS One 2013;8(1):e51926.
30. Myers RP, Tandon P, Ney M, Meeberg G, Faris P, Shaheen AA et al. Validation of the five-variable Model for End-stage Liver Disease (5vMELD) for prediction of mortality on the liver transplant waiting list. Liver Int 2013.
31. Mizock BA, Falk JL. Lactic acidosis in critical illness. Crit Care Med 1992;20(1):80-93.
32. Almenoff PL, Leavy J, Weil MH, Goldberg NB, Vega D, Rackow EC. Prolongation of the half-life of lactate after maximal exercise in patients with hepatic dysfunction. Crit Care Med 1989;17(9):870-873.
33. Olthoff KM, Kulik L, Samstein B, Kaminski M, Abecassis M, Emond J et al. Validation of a current definition of early allograft dysfunction in liver transplant recipients and analysis of risk factors. Liver Transpl 2010;16(8):943-949.
34. Powelson JA, Cosimi AB, Lewis WD, Rohrer RJ, Freeman RB, Vacanti JP et al. Hepatic retransplantation in New England--a regional experience and survival model. Transplantation 1993;55(4):802-806.
35. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29-36.
36. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004;329(7458):168-169.
37. Mirza DF, Gunson BK, Da Silva RF, Mayer AD, Buckels JA, McMaster P. Policies in Europe on "marginal quality" donor livers. Lancet 1994;344(8935):1480-1483.

38. de Vera ME, Lopez-Solis R, Dvorchik I, Campos S, Morris W, Demetris AJ et al. Liver transplantation using donation after cardiac death donors: long-term follow-up from a single center. Am J Transplant 2009;9(4):773-781.
39. Ijtsma AJ, van der Hilst CS, de Boer MT, de Jong KP, Peeters PM, Porte RJ et al. The clinical relevance of the anhepatic phase during liver transplantation. Liver Transpl 2009;15(9):1050-1055.
40. Feng L, Zhao N, Yao X, Sun X, Du L, Diao X et al. Histidine-tryptophan-ketoglutarate solution vs. University of Wisconsin solution for liver transplantation: a systematic review. Liver Transpl 2007;13(8):1125-1136.

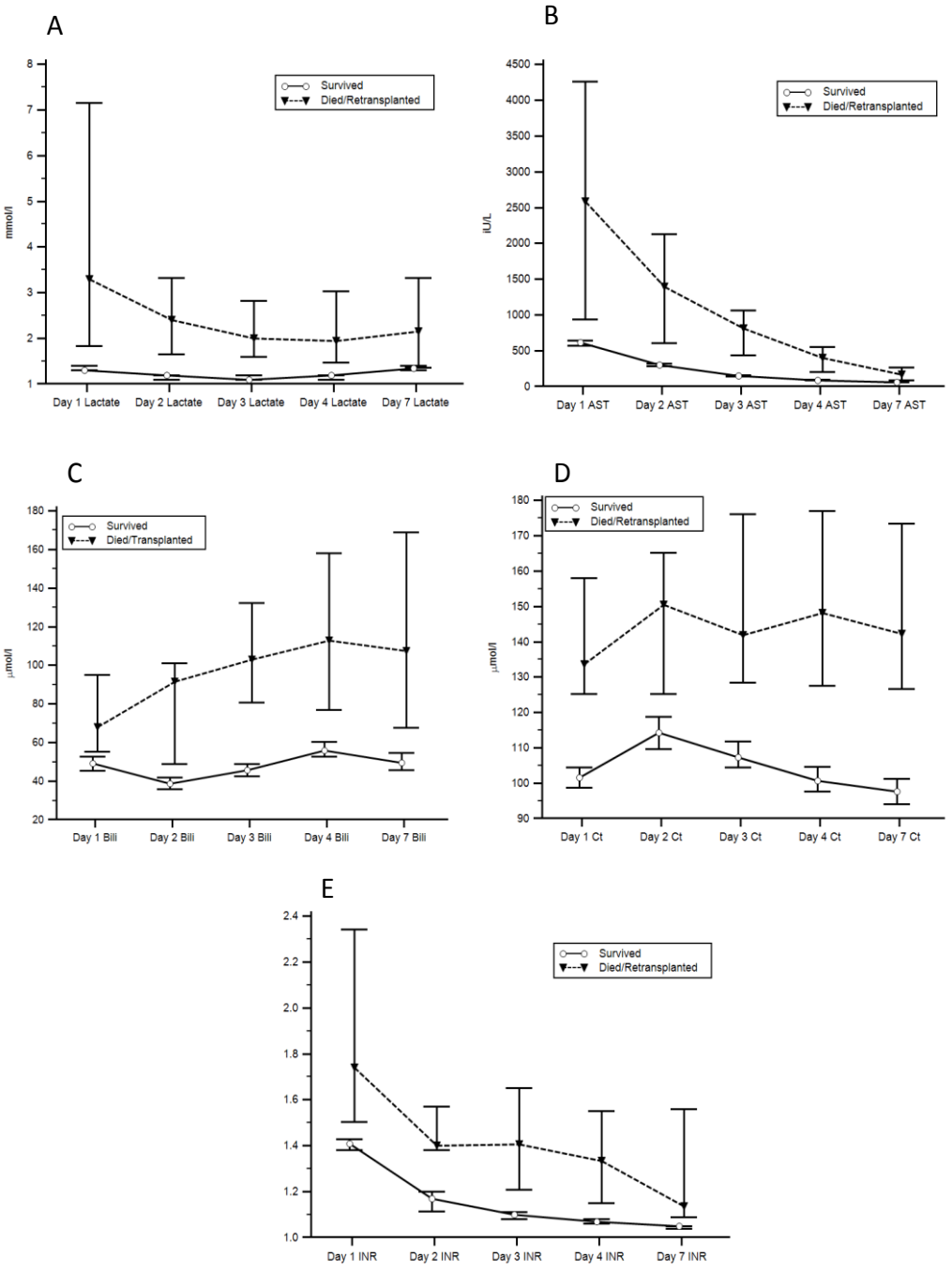

Figure 1: Sequential changes in main discriminatory biochemical indices over the first week post liver transplantation. Median (95%CI median) values are plotted. A) Lactate B) AST C) Bilirubin D) Creatinine E) INR

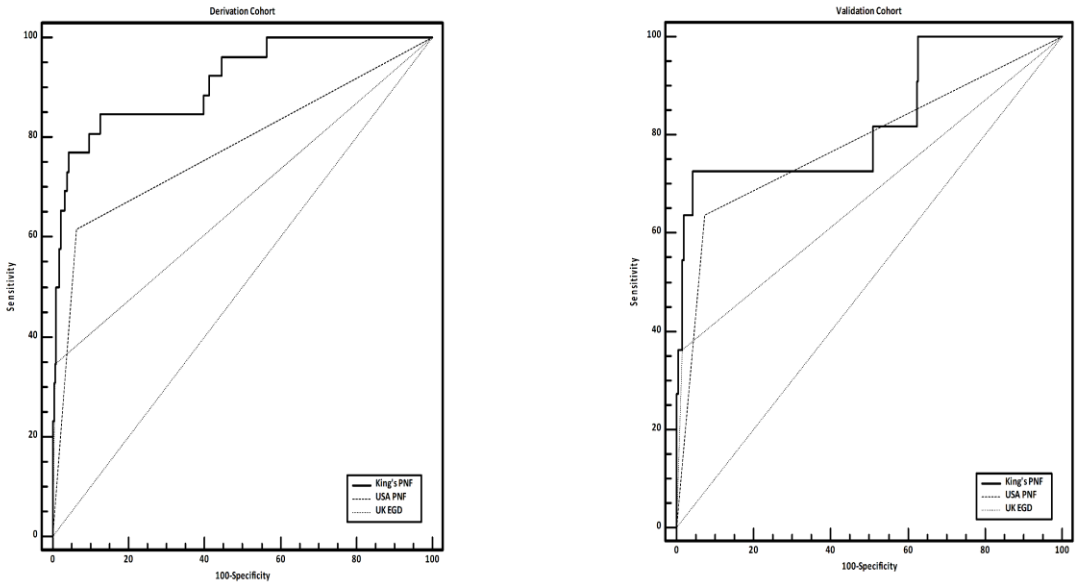

Figure 2: A) Derivation cohort comparison of AUROC curves. King's PNF AUROC 0.912 (0.889 -0.932), USA PNF AUROC 0.776(0.774-0.806, p=0.010 compared to King's PNF using Hanley & McNeil method), UK EGD AUROC 0.669 (0.634-0.704, p<0.001compared to King's PNF using Hanley & McNeil method)

B) Validation cohort comparison of AUROC curves Kings PNF AUROC 0.831 (0.789 -0.867) USA PNF 0.781(0.736-0.822, p=0.547 compared to King's PNF using Hanley & McNeil method) UK EGD 0.674 (0.624-0.721, p=0.154 compared to King's PNF using Hanley & McNeil method).

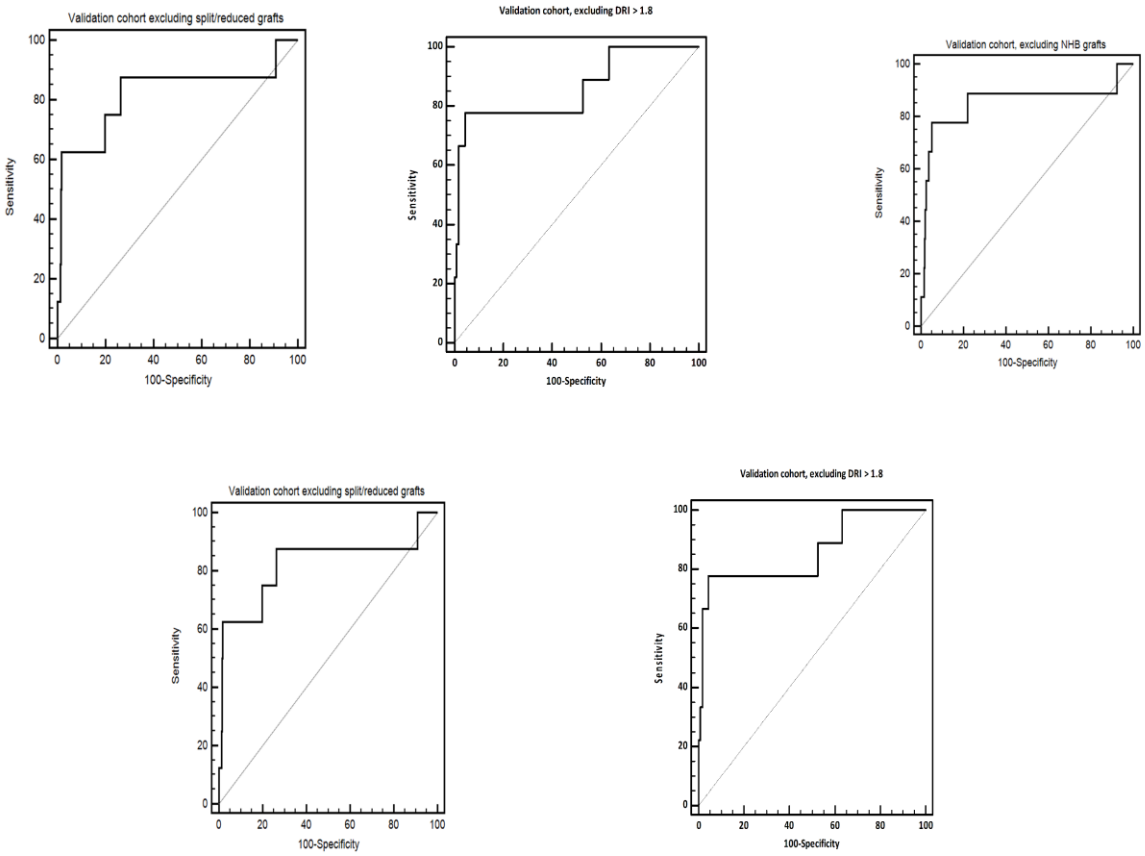

Figure 3: Comparison of the performance of King's model for diagnosis of primary non-function

a. All patients, AUROC=0.840 (0.799-0.876), p<0.001.

b. Excluding patients with ALF, AUROC=0.795 (0.746-0.839), p=0.041

c. Excluding recipients of DCD grafts, AUROC=0.854 (0.812-0.890, p<0.001

d. Excluding recipients of partial grafts, AUROC=0.821 (0.775-0.861),p=0.003

e. Excluding recipients of marginal grafts, AUROC=0.860 (0.814-0.898),p<0.001

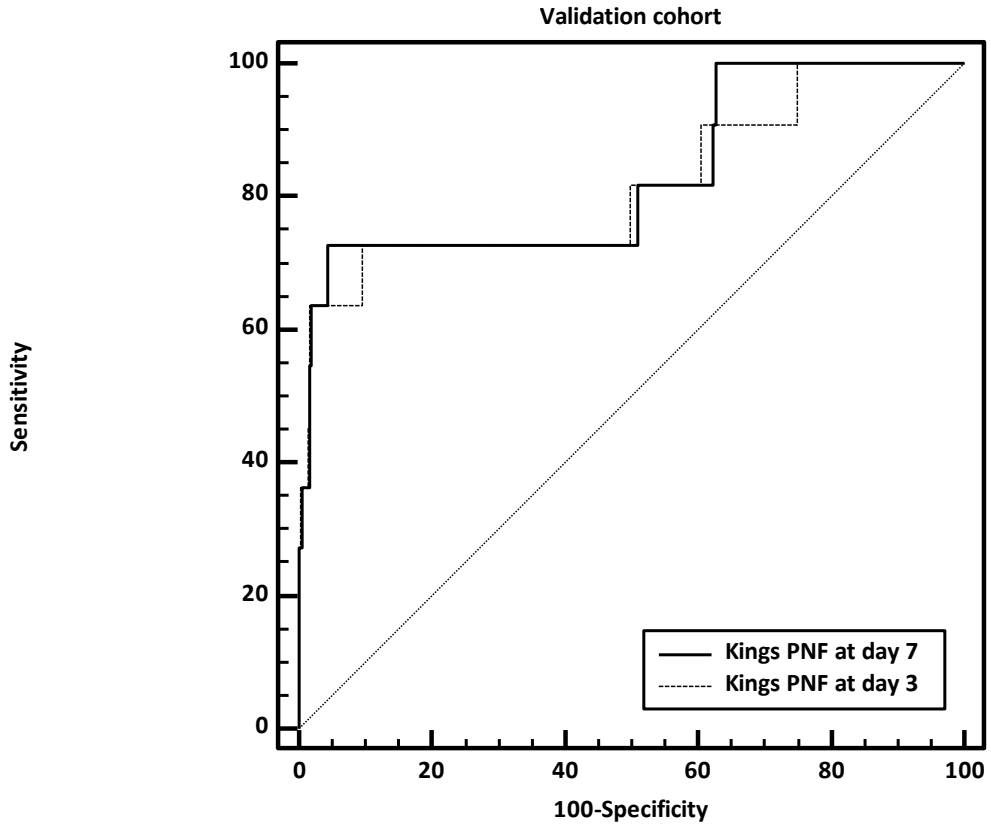

Figure 4: Pairwise comparison of AUROC for King's model of PNF inclusive and exclusive of day-7 variables (AUROC 0.831 v 0.818, p=0.321 for comparison using Hanley & McNeil method).

SUPPLEMENTARY

Supplementary methods for repeated measures logistic regression

The data were transformed to fit into the data structure required of generalised estimating equations (GEE). The outcome variable remained the dichotomous 2-week death or re-transplant variable and a binary logistic model used to allow for the dichotomous outcome variable.

All independent variables (IVs) were included in original models with those with a trajectory (AST, INR, lactate, creatinine, bilirubin, use of vasopressors) marked as repeated measures. Factors (categorical IV) included were gender, aetiology and use of advanced cardiovascular, respiratory, renal support and covariates (continuous IV) were age, AST, lactate, bilirubin, creatinine, albumin, INR, sodium and DRI. A main effects model with a random intercept was chosen.

Supplementary results for repeated measures logistic regression

The final model included AST, lactate, bilirubin and pressor requirement with highly significant predictive accuracy and a sensitivity of 71 (55-84) % and specificity of 95 (92-97) %, LR+ 14 (10-20), LR- 0.3 (0.2-0.5). This was essentially the same as with the original proposed model with an area under the ROC curve of 0.83 as detailed below.

Tests of Model Effects

	Type III		
Source	Wald Chi-Square	df	Sig.
Pressors	387.207	2	.000
AST	8.095	1	.004
Lactate	64.079	1	.000
Bilirubin	16.487	1	.000

Dependent Variable: 2WDReTx

Model: Pressors, AST, Lactate, Bilirubin

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
[Pressors=0]	-4.489	.2282	-4.937	-4.042	387.112	1	.000
[Pressors=1]	-3.330	.2846	-3.888	-2.773	136.983	1	.000
AST	.000	6.6782E-5	5.912E-5	.000	8.095	1	.004
Lactate	.341	.0426	.257	.424	64.079	1	.000
Bilirubin	.006	.0015	.003	.009	16.487	1	.000
(Scale)	1						

Dependent Variable: 2WDReTx

Model: Pressors, AST, Lactate, Bilirubin

It is interesting that using a more formal approach to repeated measures without significant supervised variable selection led to a model with similar prognostic accuracy. The model we propose is a relatively simple equation and unless these more data driven techniques can provide much higher accuracy we do not recommend them at present. Also, they require data to day 7 and our model can allow stratification at day 3 where prompt preparations for re-transplantation can be started. We would agree that these generalised models are useful to explore further but suggest they could be part of a future manuscript on their use in these situations. However they do explicitly use bilirubin as part of their definition so they are incorporating the trajectory of this parameter which was not done in our simpler model.